

AD-A051 863

INDIANA UNIV BLOOMINGTON DEPT OF MATHEMATICS
ASYMPTOTIC MULTINORMALITY AND REMAINDER TERMS OF LINEAR RANK VE--ETC(U)
FEB 78 M L PURI, N S RAJARAM

AFOSR-76-2927

F/G 12/1

NL

UNCLASSIFIED

| OF |
AD
A051863

AFOSR-TR-78-0434

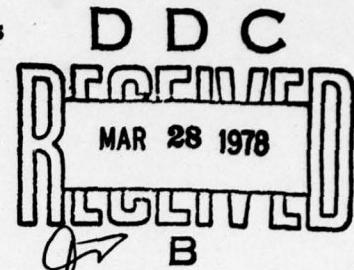


END
DATE
FILED

4 - 78

DDC

AD A051863

AD No.
DDC FILE COPYAsymptotic Multinormality and Remainder Terms
of Linear Rank Vectors Under Alternativesby
Madan L. Puri and Navaratna S. Rajaram
Indiana University and Kent State University0. Summary

Asymptotic multinormality of linear rank statistics based on independent vector valued random variables is obtained. Under suitable assumptions, weak estimates for the remainder terms for convergence to normality are also obtained. Results on asymptotic normality are related to Hájek (1968) and Puri and Sen (1969). Results on the remainder terms are related to those of Jurečková and Puri (1975), Bergström and Puri (1977), and Hušková (1977).

1. Preliminaries

Let $\underline{x}_{Ni} = (x_{Ni}^{(1)}, \dots, x_{Ni}^{(p)})'$, $i=1, \dots, N$ be a sequence of independent p -variate ($p \geq 1$) random vectors having continuous cumulative distribution functions (cdf) $F_{Ni}(\underline{x})$,

$\underline{x} = (x^{(1)}, \dots, x^{(p)})'$, $i=1, \dots, N$ respectively. Consider now the random matrix \underline{X}_N corresponding to $(\underline{x}_{N1}, \dots, \underline{x}_{NN})$, i.e.,

$$(1.1) \quad \underline{X}_N = ((x_{Ni}^{(v)}))_{\substack{i=1, \dots, N; \\ v=1, \dots, p.}} \quad ; \quad \underline{x}_{Ni} = (x_{Ni}^{(1)}, \dots, x_{Ni}^{(p)})'$$

DISTRIBUTION STATEMENT A
Approved for public release; Distribution Unlimited

Work supported by the Air Force Office of Scientific Research, AFS USAF, Grant No. AFOSR 76-2927. Reproduction in whole or part permitted for any purpose of the United States Government.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC

This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b). Distribution is unlimited.

A. D. BLOSE

Technical Information Officer

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

18 19 REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM	
REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER	
AFOSR TR-78-0434			
4. TITLE (and Subtitle) (6) ASYMPTOTIC MULTINORMALITY AND REMAINDER TERMS OF LINEAR RANK VECTORS UNDER ALTERNATIVES.		5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.	
7. AUTHOR(s) (10) Madan L. Puri - Navaratna S. Rajaram		6. PERFORMING ORG. REPORT NUMBER (15) AFOSR-76-2927	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Indiana University Department of Mathematics✓ Bloomington, Indiana 47401		10. PROGRAM ELEMENT, PROJECT, AREA & WORK UNIT NUMBERS 611021 16 2304 A5	
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE February 28, 1978	
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		13. NUMBER OF PAGES 25	
(11) 28 Feb 78		(12) 27P.	
15. SECURITY CLASS. (of this report) UNCLASSIFIED			
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) multinormality, scores, C_r-inequality, Liapunov condition, eigen-value, euclidean norm.			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Asymptotic multinormality of linear rank statistics based on independent vector valued random variables is obtained. Under suitable assumptions, weak estimates for the remainder terms for convergence to normality are related to Hájek (Ann. Math. Stat. (1968), 325-346). Results on the remainder terms are related to those of Jurečkova and Puri (Ann. Prob. (1975), 526-533), and Bergström and Puri (Ann. Stat. (1977), 671-680). compared.			

and observe that each row of \underline{x}_N is composed of N independent univariate random variables. Let $R_{Ni}^{(v)}$ be the rank of $x_{Ni}^{(v)}$ among $(x_{N1}^{(v)}, \dots, x_{NN}^{(v)})$ for each $v = 1, \dots, p$. Then corresponding to the observation matrix \underline{x}_N , we have a rank collection matrix \underline{R}_N , where

$$(1.2) \quad \underline{R}_N = ((R_{Ni}^{(v)}))_{\substack{i=1, \dots, N; \\ v=1, \dots, p.}}$$

Consider now p sets of scores $(a_{Ni}^{(v)}, 1 \leq i \leq N), 1 \leq v \leq p$ generated by known functions $\varphi_v: (0,1) \rightarrow R$ in either of the following ways:

$$(1.3) \quad a_{Ni}^{(v)} = \varphi_v(i/(N+1), i=1, \dots, N; v=1, \dots, p)$$

$$(1.4) \quad a_{Ni}^{(v)} = E\varphi_v(U_N^{(i)}), i=1, \dots, N; v=1, \dots, p$$

where $U_N^{(i)}$ is the i^{th} order statistic in a random sample of size N from the uniform distribution over (0,1). Consider now the simple linear rank vector \underline{s}_N corresponding to \underline{x}_N

(or \underline{R}_N) where

$$(1.5) \quad \underline{s}_N = (s_N^{(1)}, \dots, s_N^{(p)}), s_N^{(v)} = \sum_{i=1}^N c_{Ni}^{(v)} a_N^{(v)} (R_{Ni}^{(v)}), 1 \leq v \leq p$$

and where $(c_{Ni}^{(v)}, 1 \leq i \leq N), 1 \leq v \leq p$ are p-sets of known (regression) constants.

NTIS	White Section
DDC	Buff Section <input type="checkbox"/>
UNANNOUNCED	
JUSTIFICATION _____	
BY _____	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
A	

In the next section we establish the asymptotic normality of \tilde{S}_N by following the methods of Chernoff-Savage (1958) and Puri-Sen (1971). For a different approach, see Hájek (1968) and Puri-Sen (1969). In Section 3 we obtain the estimates for the order of normal approximation for \tilde{S}_N . This constitutes the generalizations of the results of Jurečková-Puri (1975), Bergstrom-Puri (1977), Puri-Rajaram (1977), and Hušková (1977b), where the problems are treated in the univariate set-ups. The multivariate extensions in the generality of our paper do not appear to exist in the literature so far. (For a rather special multivariate case, the reader is referred to Hušková (1977a)).

2. Asymptotic Normality of \tilde{S}_N . We now establish the asymptotic normality of \tilde{S}_N defined in (1.5). We make the following assumptions:

$$(2.1) \quad |\phi_v(t)| \leq K\{(1-t)\}^{\delta - \frac{1}{2}}, \quad |\phi'_v(t)| \leq K\{t(1-t)\}^{\delta - \frac{3}{2}}; \quad 0 < t < 1$$

$$(2.2) \quad \max_{1 \leq i \leq N} |c_{Ni}^{(v)}| \frac{s_v^{-1}}{s_v} = O(N^{-\frac{1}{2}}), \quad 1 \leq v \leq p$$

where

$$(2.3) \quad s_v^2 = \text{var } S_N^{(v)}, \quad 1 \leq v \leq p$$

For convenience we shall take $\tilde{S}_N = (\frac{s_N^{(1)}}{s_1}, \dots, \frac{s_N^{(p)}}{s_p})'$, where the s_v 's are given by (2.3).

We then have the following theorem.

Theorem 2.1 Let the scores $a_{Ni}^{(v)}$, $1 \leq v \leq p$ be defined by (1.3).

Then, under the assumptions (2.1) and (2.2), for every vector

λ in R^p , $\lambda'(\tilde{S}_N - u_N)/(\lambda' v_N \lambda)^{\frac{1}{2}}$ has asymptotically (as $N \rightarrow \infty$)

the $n(0,1)$ distribution where u_N and v_N are defined by

$$(2.4) \quad \underline{\mu}_N = (\mu_N^{(1)}, \dots, \mu_N^{(p)})', \quad \mu_N^{(v)} = \sum_{i=1}^N C_{Ni}^{(v)} \int_{-\infty}^{\infty} \varphi_v(H^{(v)}(x)) dF_{Ni}^{(v)}(x), \quad 1 \leq v \leq p$$

$$(2.5) \quad H^{(v)}(x) = N^{-1} \sum_{i=1}^N F_{Ni}^{(v)}(x); \quad F_{Ni}^{(v)}(x) = P(X_{Ni}^{(v)} \leq x), \quad 1 \leq v \leq p$$

$$(2.6) \quad \underline{v}_N = \left(\left(\frac{s_{\mu v}}{s_{\mu} s_v} \right) \right), \quad 1 \leq \mu, v \leq p$$

and

$$(2.7) \quad s_{\mu v} = \text{Cov}(S_N^{(\mu)}, S_N^{(v)}); \quad 1 \leq \mu, v \leq p.$$

Furthermore, the theorem remains true, if s_{μ} in (2.7)

are given by

$$(2.8) \quad s_{\mu v} = \sum_{i=1}^N \text{Cov}(A_{Ni}^{(\mu)}(X_{Ni}^{(\mu)}), A_{Ni}^{(v)}(X_{Ni}^{(v)})),$$

where

$$(2.9) \quad A_{Ni}^{(v)}(x) = \frac{1}{N} \sum_{j=1}^N (C_{Nj}^{(v)} - C_{Ni}^{(v)}) \int_{-\infty}^{\infty} \{I_{[x \leq y]} - F_{Ni}^{(v)}(y)\} \varphi'_v(H_{(j)}^{(v)}(y)) dF_{Nj}^{(v)}(y)$$

and

$I(x \leq y) = 1$ if $x \leq y$, and 0 otherwise

Proof:

Let $(\alpha_1, \alpha_2, \dots, \alpha_p)$ be a set of fixed but arbitrary constants.

By the Cramer-Wold criterion, it suffices to show that

$$(2.10) \quad \sum_{v=1}^p \alpha_v \frac{s_N^{(v)}}{s_{\mu}} = U_N, \quad \text{say is asymptotically normal.}$$

We introduce the following representation for $S_N^{(v)}$.

Let

$$(2.11) \quad H_N^{(v)}(x) = N^{-1} \sum_{i=1}^N I(X_{Ni}^{(v)} \leq x),$$

$H^{(v)}(x)$ is already defined. (See (2.4) above.)

$$(2.12) \quad C_N^{(v)}(x) = \sum_{i=1}^N C_{Ni}^{(v)} I(X_{Ni}^{(v)} \leq (x))$$

$$(2.13) \quad C^{(v)}(x) = \sum_{i=1}^N C_{Ni}^{(v)} F_{Ni}^{(v)}(x) = E[C_N^{(v)}(x)]$$

We shall adhere to the convention of denoting stochastic variables

$(H_N^{(v)}, C_N^{(v)})$ with subscripts and non-random functions $(H^{(v)}, C^{(v)})$

without subscripts, although depending upon N .

Then the following inequalities are immediate:

$$(2.14) \quad |C_N^{(v)}(x)| \leq N \max_{1 \leq i \leq N} |C_{Ni}^{(v)}| H_N^{(v)}(x)$$

$$(2.15) \quad |C^{(v)}(x)| \leq N \max_{1 \leq i \leq n} |C_{Ni}^{(v)}| H^{(v)}(x), \quad 1 \leq v \leq p$$

We shall use the representation.

$$(2.16) \quad S_N^{(v)} = \int_{-\infty}^{\infty} \varphi_v\left(\frac{N}{N+1} H_N^{(v)}(x)\right) dC_N^{(v)}(x) = \mu_N^{(v)} + \beta_{1N}^{(v)} + \beta_{2N}^{(v)} + \sum_{j=1}^3 D_{jN}^{(v)}$$

where $\mu_N^{(v)}$ is given by (2.3).

$$(2.17) \quad \beta_{1N}^{(v)} = \int_{-\infty}^{\infty} \varphi_v(H_N^{(v)}(x)) d(C_N^{(v)}(x) - C^{(v)}(x))$$

$$(2.18) \quad \beta_{2N}^{(v)} = \int_{-\infty}^{\infty} (H_N^{(v)}(x) - H^{(v)}(x)) \varphi'_v(H^{(v)}(x)) dC^{(v)}(x)$$

$$(2.19) \quad D_{1N}^{(v)} = \frac{-1}{N+1} \int_{-\infty}^{\infty} H_N^{(v)}(x) \varphi'_v(H^{(v)}(x)) dC_N^{(v)}(x)$$

$$(2.20) \quad D_{2N}^{(v)} = \int_{-\infty}^{\infty} (H_N^{(v)}(x) - H^{(v)}(x)) \varphi_v(H^{(v)}(x)) d(C_N^{(v)}(x) - C^{(v)}(x))$$

$$(2.21) \quad D_{3N}^{(v)} = \int_{-\infty}^{\infty} \left\{ \varphi_v\left(\frac{N}{N+1} H_N^{(v)}(x)\right) - \varphi_v(H^{(v)}(x)) - \left(\frac{N}{N+1} H_N^{(v)}(x) - H^{(v)}(x)\right) \right\} dC_N^{(v)}(x)$$

Substituting (2.16) in (2.10), we have

$$(2.22) \quad u_N = \sum_{v=1}^p \frac{a_v}{g_v} \left\{ u_N^{(v)} + (\beta_{1N}^{(v)} + \beta_{2N}^{(v)}) + \sum_{j=1}^3 D_{jN}^{(v)} \right\}$$

The proof will be accomplished if we show the following:

(A) $\mu_N^{(v)}$ is finite.

(B) $\sum_{v=1}^p \frac{a_v}{g_v} (\beta_{1N}^{(v)} + \beta_{2N}^{(v)})$ is asymptotically normal.

(C) $\sum_{v=1}^p \frac{a_v}{g_v} \sum_{j=1}^3 D_{jN}^{(v)} \rightarrow 0$ in probability.

Proof of (A): Observe that

$$|\mu_N^{(v)}| = \left| \int_{-\infty}^{\infty} \varphi_v(H^{(v)}(x)) dC^{(v)}(x) \right|$$

$$\leq N \max_{1 \leq i \leq N} |C_{Ni}^{(v)}| \int_{-\infty}^{\infty} \varphi_v(H^{(v)}(x)) dH^{(v)}(x) < +\infty \text{ for each } N$$

by assumptions (2.1) and (2.15).

Proof of (B): For a fixed v , we shall verify the Liapunov criterion for the normality of $(\beta_{1N}^{(v)} + \beta_{2N}^{(v)}) g_v^{-1}$ and then do the same

for the sum $\sum_{v=1}^p a_v (\beta_{1N}^{(v)} + \beta_{2N}^{(v)}) g_v^{-1}$, by an extension of the C_r -inequality.

In fact we do so separately for $\beta_{1N}^{(v)} g_v^{-1}$ and $\beta_{2N}^{(v)} g_v^{-1}$ and use the C_r -inequality.

Since we are considering a fixed v , we can drop the indexing variable v and simplify our notation substantially.

Consider (2.18)

$$\beta_{2N} = \int_{-\infty}^{\infty} (H_N(x) - H(x)) \varphi'(H(x)) dC(x)$$

Integration by parts yields,

$$(2.23) \quad \beta_{2N} = B(x) \{H_N(x) - H(x)\} \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{\infty} B(x) d(H_N(x) - H(x))$$

where

$$(2.24) \quad B(x) = \int_{x_0}^x \varphi'(H(x)) dC(x), \quad x_0 \text{ arbitrary but } H(x_0) > 0.$$

We shall later prove that the first term in (2.23) is

$$O_p(g), \quad g_v = g.$$

The second term of the same can be written as

$$(2.25) \quad \int_{-\infty}^{\infty} B(x) d(H_N(x) - H(x)) N^{-1} \sum_{i=1}^N \{B(x_{Ni}) - E\beta(x_{Ni})\}$$

We shall show that

$$(2.26) \quad (\frac{1}{N})^{(2+\epsilon)} \sum_{i=1}^N E|\beta(x_{Ni}) - E\beta(x_{Ni})|^{2+\epsilon} \rightarrow 0, \quad \text{as } N \rightarrow \infty \text{ for some } \epsilon > 0.$$

In fact, it suffices to verify,

$$(2.27) \quad (\frac{1}{N})^{(2+\epsilon)} \sum_{i=1}^N E|\beta(x_{Ni})|^{2+\epsilon} \rightarrow 0$$

which by the C_r and Jensen inequalities implies (2.26).

Next, choose an $\epsilon > 0$ such that $(2+\epsilon)(\delta - \frac{1}{2}) > -1$ for the δ given by the constraints on φ . Then,

$$(2.28) (\bar{s}_N)^{(2+\epsilon)} \sum_{i=1}^N |\beta(x_{Ni})|^{2+\epsilon} = (\bar{s}_N)^{(2+\epsilon)} \sum_{i=1}^N E|\beta(x_{Ni})|^{2+\epsilon}$$

$$\begin{aligned} &\leq (\bar{s}_N)^{(2+\epsilon)} \sum_{i=1}^N N^{2+\epsilon} \max_{1 \leq i \leq N} |c_{Ni}|^{2+\epsilon} \int_{-\infty}^{\infty} \left| \int_{x_0}^x \varphi'(H(y)) dH(y) \right| dF_{Ni}(x) \\ &= \left\{ \max_{1 \leq i \leq N} \frac{1}{s} |c_{Ni}| \right\} \sum_{i=1}^N \int_{-\infty}^{\infty} \left\{ |\varphi(H(x))| + |\varphi(H(x_0))| \right\}^{2+\epsilon} dF_{Ni}(x) \\ &= O(N^{-\epsilon/2}) \int_{-\infty}^{\infty} \left\{ |\varphi(H(x))| + |\varphi(H(x_0))| \right\}^{2+\epsilon} dH(x) \rightarrow 0, \text{ as } N \rightarrow \infty. \end{aligned}$$

because $(2+\epsilon)(\delta - \frac{1}{2}) > -1$ and $\varphi(H(x_0))$ is a constant. Observe

that we have used the fact $\max_{1 \leq i \leq N} \frac{|c_{Ni}|}{s^{2+\epsilon}} = O(N^{-1-\epsilon/2})$.

It remains to be shown that in (2.23),

$$\left| \frac{\beta(x)}{s} \right|_{-\infty}^{\infty} = \frac{B(x)}{s} \left\{ H_N(x) - H(x) \right\} \Big|_{-\infty}^{+\infty} = o_p(1).$$

$$\begin{aligned} \left| \frac{\beta(x)}{s} \right| &= \sqrt{N} |H_N(x) - H(x)| \frac{1}{\sqrt{Ns}} \left| \int_{x_0}^x \varphi'(H(y)) dH(y) \right| \\ &\leq \sqrt{N} |H_N(x) - H(x)| O(1) \left| \int_{x_0}^x \varphi'(H(y)) dH(y) \right| \end{aligned}$$

$$\leq K \sqrt{N} |H_N(x) - H(x)| \left\{ H(x)(1 - H(x)) \right\}^{\delta - \frac{1}{2}}$$

since $\varphi(H(x_0))$ is a constant.

By Puri-Sen (1971), given any ϵ, δ' positive, there is a constant $C(\epsilon, \delta')$ independent of N such that, with probability $> 1-\epsilon$,

$$\frac{1}{N^2} |H_N(x) - H(x)| \leq C(\epsilon, \delta') \left\{ H(x)(1-H(x)) \right\}^{\frac{1}{2}-\delta'}.$$

Thus,

$$|\bar{s}_N^{-1} \beta(x)| \leq K \left\{ H(x)(1-H(x)) \right\}^{\delta-\delta'} C(\epsilon, \delta') \rightarrow 0$$

as $x \rightarrow \pm\infty$, by choosing $0 < \delta' < \delta$.

The verification of the Liapunov condition for $\bar{s}_{1N}^{-1} \beta_{1N}$ is similar, in fact, easier and is therefore not given here. The

C_r -inequality yields the Liapunov condition for $\bar{s}_r^{-1} (\beta_{1N} + \beta_{2N})$

Consider next the normalized sum,

$$\sum_{v=1}^p \alpha_v \bar{s}_v^{-1} (\beta_{1N}^{(v)} + \beta_{2N}^{(v)}) .$$

We need to verify the Liapunov condition for the above expression.

$$\text{Let } \sigma^2 = \text{Var} \left(\sum_{v=1}^p \alpha_v \bar{s}_v^{-1} (\beta_{1N}^{(v)} + \beta_{2N}^{(v)}) \right) .$$

We shall assume that σ^2 is bounded away from zero for all N .
(If not, the sum is trivially degenerate normal as $N \rightarrow +\infty$.)

$$\text{Write } \beta_{1N}^{(v)} + \beta_{2N}^{(v)} = \sum_{i=1}^N \beta_{Ni}^{(v)}$$

where $\beta_{Ni}^{(\nu)}$ are independent random variables (as already done earlier). We have to show

$$(2.29) \quad \lim_{N \rightarrow \infty} \sigma^{-(2+\epsilon)} \sum_{i=1}^N E \left| \sum_{\nu=1}^p \alpha_\nu \frac{\beta_{Ni}^{(\nu)}}{g_\nu} \right|^{2+\epsilon} = 0 .$$

We have already shown that

$$(2.30) \quad \lim_{N \rightarrow \infty} (g_\nu)^{(2+\epsilon)} \sum_{i=1}^N E \left| \beta_{Ni}^{(\nu)} \right|^{2+\epsilon} = 0$$

(2.29) follows from (2.30) upon noting that there is a constant $C(p, \epsilon)$, depending only upon p and ϵ such that

$$E \left| \sum_{\nu=1}^p \alpha_\nu \frac{\beta_{Ni}^{(\nu)}}{g_\nu} \right|^{2+\epsilon} \leq C(p, \epsilon) \sum_{\nu=1}^p |\alpha_\nu / g_\nu|^{2+\epsilon} E \left| \beta_{Ni}^{(\nu)} \right|^{2+\epsilon}$$

(Generalized C_r -inequality.)

This establishes (B). We note that ϵ depends only upon δ and hence the same choice of ϵ works for all ν .

Proof of (C): Recall that we have to establish

$$\sum_{\nu=1}^p \alpha_\nu g_\nu^{-1} \sum_{j=1}^3 D_{jN}^{(\nu)} = o_p(1) .$$

Clearly, it suffices to prove for a particular ν since we have a finite sum $(1, \dots, p)$. Again, we shall drop the index ν .

Consider, $|\bar{g}^l D_{1N}| = |-(N+1)^{-1} g^{-1}| \int_{-\infty}^{\infty} H_N(x) \varphi'(H(x)) dC_N(x) |$

$$\leq (gN)^{-1} \sum |\varphi'(H(x_{Ni}))| |C_{Ni}| = \frac{1}{N} \sum_{i=1}^N v_{Ni}$$

where

$$v_{Ni} = |\varphi'(H(x_{Ni})) \bar{s}^{-1} c_{Ni}| .$$

To prove

$$\sum_{i=1}^N v_{Ni} \rightarrow 0 \text{ in probability},$$

it suffices to show that

$$\sum_{i=1}^N \bar{N}^\alpha E|v_{Ni}|^\alpha < +\infty, \text{ for all } N, \text{ for some } \alpha, 0 < \alpha < 1 .$$

(By particular case 1° , page 241, Loéve (1963)).

Take $\alpha = 2/3$:

$$\begin{aligned} N^{-2/3} \sum_{i=1}^N E|v_{Ni}|^{2/3} &\leq K N^{-2/3} \sum_{i=1}^N \left| \frac{c_{Ni}}{\bar{s}} \right|^{2/3} E|H(x_{Ni})(1 - H(x_{Ni}))|^{(\delta-3/2)} \\ &\leq N^{1/3} \max_{1 \leq i \leq N} (\bar{s}^{-1} |c_{Ni}|)^{2/3} \sum_{i=1}^N \int_{-\infty}^{\infty} \{H(x)(1 - H(x))\}^{\frac{2\delta}{3}-1} dF_{Ni}(x) \\ &= K \cdot O(1) \int_{-\infty}^{\infty} \{H(x)(1 - H(x))\}^{\frac{2\delta}{3}-1} dH(x) < +\infty ; \end{aligned}$$

where we have used the fact that

$$\max_{1 \leq i \leq N} |\bar{s}^{-1} c_{Ni}|^{2/3} = O(N^{-1/3}) .$$

Consider next,

$$D_{2N} = \int_{-\infty}^{\infty} (H_N(x) - H(x)) \varphi'(H(x)) d(C_N(x) - C(x)) .$$

By Puri-Sen (1971), given any $\epsilon > 0$, $0 < \delta' < \frac{1}{2}$, there is a constant $C(\epsilon, \delta')$, independent of N such that with probability $> 1 - \epsilon$,

$$N^{\frac{1}{2}} |H_N(x) - H(x)| < C(\epsilon, \delta') \left\{ H(x)(1 - H(x)) \right\}^{\frac{1}{2} - \delta'}.$$

Thus with probability $> 1 - \epsilon$,

$$|H_N(x) - H(x)| |\psi'(H(x))| \leq KC(\epsilon, \delta') N^{\frac{1}{2}} \{H(x)(1 - H(x))\}^{\delta^*-1}$$

taking $\delta' < \delta$ and setting $\delta^* = \delta - \delta'$.

It suffices to show that

$$\bar{s}^1 \bar{N}^{\frac{1}{2}} \int_{-\infty}^{\infty} \{H(x)(1 - H(x))\}^{\delta^*-1} d(C_N(x) - C(x)) \rightarrow 0 \text{ in probability.}$$

We shall use the Liapunov criterion for degenerate convergence (page 275, Loeve (1963)).

$$\begin{aligned} & \bar{s}^1 \bar{N}^{\frac{1}{2}} \int_{-\infty}^{\infty} \{H(x)(1 - H(x))\}^{\delta^*-1} dC_N(x) \\ &= \bar{s}^1 \bar{N}^{\frac{1}{2}} \sum_{i=1}^N C_{Ni} \{H(X_{Ni})(1 - H(X_{Ni}))\}^{\delta^*-1} \end{aligned}$$

Set

$$v_{Ni} = N^{\frac{1}{2}} \bar{s}^{-1} C_{Ni} \{H(X_{Ni})(1 - H(X_{Ni}))\}^{\delta^*-1}.$$

Then,

$$\bar{s}^1 \bar{N}^{\frac{1}{2}} \sum_{i=1}^N C_{Ni} \{H(X_{Ni})(1 - H(X_{Ni}))\}^{\delta^*-1} = \bar{N}^{-1} \sum_{i=1}^N v_{Ni}.$$

It remains to prove, $\bar{N}^{-1} \sum_{i=1}^N (v_{Ni} - E v_{Ni}) \rightarrow 0$ in probability.

This will be accomplished if we can show that for some $\alpha > 0$,

$$(2.31) \bar{N}^{(1+\alpha)} \sum_{i=1}^N E |v_{Ni}|^{1+\alpha} \rightarrow 0 .$$

Choose $\alpha > 0$ such that, $(1+\alpha)(\delta^* - 1) > -1$; that is,
 $0 < \alpha < \frac{\delta^*}{1-\delta^*}$. Then,

$$\begin{aligned} & \bar{N}^{(1+\alpha)} \sum_{i=1}^N E |v_{Ni}|^{1+\alpha} \\ & \leq \bar{N}^{1(1+\alpha)} \left| \frac{\sqrt{N} \max_{1 \leq i \leq N} |C_{Ni}|}{s} \right|^{1+\alpha} \left| \sum_{i=1}^N E \{H(X_{Ni}) (1 - H(X_{Ni}))\}^{(1+\alpha)(\delta^*-1)} \right. \\ & = O(1) \bar{N}^\alpha \int_{-\infty}^{\infty} \{H(x) (1 - H(x))\}^{(1+\alpha)(\delta^*-1)} dH(x) = O(N^{-\alpha}) \end{aligned}$$

which proves (2.31).

Next consider

$$\bar{s}^{\frac{1}{2}} D_{3N} = \bar{s}^{\frac{1}{2}} \int_{-\infty}^{\infty} \left\{ \varphi \left(\frac{NH_N(x)}{N+1} \right) - \varphi(H(x)) - \left(\frac{NH_N(x)}{N+1} - H(x) \right) \varphi'(H(x)) \right\} dC_N(x) .$$

We have to show that $\bar{s}^{\frac{1}{2}} D_{3N} \rightarrow 0$ in probability.

The following substitution will simplify the proof. Let

$$(2.32) C_{3N} = \frac{D_{3N}}{s \sqrt{N}} .$$

Then it suffices to prove $C_{3N} = O_p(N^{-\frac{1}{2}})$

Observe that, since $\max_{1 \leq i \leq N} \left| \frac{C_{Ni}}{\sqrt{N}s} \right| = O(N^{-1})$

we have

$$(2.33) \quad |c_{3N}|$$

$$\leq O(1) \int_{-\infty}^{\infty} |\varphi\left(\frac{N}{N+1} H_N(x)\right) - \varphi(H(x)) - \left(\frac{N}{N+1} - H(x)\right) \varphi'(H(x))| dH_N(x)$$

The proof that the right hand side of (2.36) is $O_p(N^{-\frac{1}{2}})$ can be found in Puri-Sen (1971), pages 401-405.

This proves Theorem 2.1.

For a different approach, as well as the derivation of some tests for linear hypothesis based on S_N , see Puri-Sen (1969).

Computation of V_N : It is clear that the variance contribution is from the term $\beta_{1N}^{(v)} + \beta_{2N}^{(v)}$. It is easy to check that

$$(2.34) \quad \beta_{1N}^{(v)} + \beta_{2N}^{(v)} = \sum_{i=1}^N \left\{ A'_{Ni}(x^{(v)}) \left(x^{(v)}_{Ni} \right) - E A'_{Ni}(x^{(v)}_{Ni}) \right\}$$

where

$$A'_{Ni}(x) = \frac{1}{N} \sum_{j=1}^N (C_{Ni}^{(v)} - C_{Nj}^{(v)}) \int_{x_0}^x \varphi'_v(H^{(v)}(y)) dF_{Nj}^{(v)}(y); \quad x_0 \text{ arbitrary.}$$

Centering $A'_{Ni}(x)$ at its expectations, we note that

$$(2.35) \quad \beta_{1N}^{(v)} + \beta_{2N}^{(v)} = \sum_{i=1}^N A'_{Ni}(x^{(v)}_{Ni})$$

where

$$A_{Ni}^{(v)}(x) = \frac{1}{N} \sum_{j=1}^N (C_{Ni}^{(v)} - C_{Nj}^{(v)}) \int_{-\infty}^{\infty} \{I(x \leq y) - F_{Ni}^{(v)}(y)\} \varphi_v'(H^{(v)}(y)) dF_{Nj}^{(v)}(y)$$

which yields the desired approximation (2.7).

We next give a corollary which extends these results to the case when the scores are generated according to (1.7)(b).

Corollary 2.2: Let φ_v be the inverse of a distribution function. Let s_N^* be the linear rank vector defined as in (1.8) but with the scores generated according to (1.4). Then, under the conditions of Theorem 2.1, s_N^* is asymptotically multinormal with the same parameters μ_N and V_N .

Proof: Let

$$(2.36) \quad s_N^* = \left(\frac{s_N^{*(1)}}{g_1}, \dots, \frac{s_N^{*(p)}}{g_p} \right)^T.$$

Clearly, it suffices to prove

$$(2.37) \quad s_N^{*(v)} - s_N^{(v)} = o_p(g_v), \quad 1 \leq v \leq p$$

which then entails $s_N^* - s_N \rightarrow 0$ in probability.

Again, since v is fixed, $1 \leq v \leq p$, we shall drop this index.

We define

$$(2.38) \quad \varphi_N(t) = \sum_{i=1}^{[Nt]} (G_N(i) - G_N(i-1)), \quad G_N(0) = 0, \quad 0 < t < 1,$$

where $[a]$ is the greatest integer not exceeding a .

It is easy to check that

$$(2.39) \quad s_N^* = \int_{-\infty}^{\infty} \varphi_N\left(\frac{N H_N(x)}{N+1}\right) d C_N(x) .$$

The proof of Corollary 2.2 will be accomplished if we prove the following:

Lemma 2.3: Under the hypotheses of Theorem 2.1 and Corollary 2.2, we have

$$(i) \quad \lim_{N \rightarrow +\infty} \varphi_N(t) = \varphi(t)$$

$$(ii) \quad \left| \int_{-\infty}^{\infty} \left\{ \varphi_N\left(\frac{N}{N+1} H_N(x)\right) - \varphi\left(\frac{N}{N+1} H_N(x)\right) \right\} d C_N(x) \right| = o_p(g)$$

Proof of Lemma 2.3: We shall only prove (ii) since the proof of (i) can be found in pages 408-409 of Puri-Sen (1971).

By inequality (2.14), we have

$$(2.40) \quad \left| \int_{-\infty}^{\infty} \left\{ \varphi_N\left(\frac{N H_N(x)}{N+1}\right) - \varphi\left(\frac{N H_N(x)}{N+1}\right) \right\} d C_N(x) \right|$$

$$\leq \max_{1 \leq i \leq N} |C_{Ni}| \sum_{i=1}^N \left| \varphi_N\left(\frac{i}{N+1}\right) - \varphi\left(\frac{i}{N+1}\right) \right| .$$

Hence it follows that

$$(2.41) \quad \begin{aligned} & \max_{1 \leq i \leq N} \left| \frac{C_{Ni}}{g} \right| \sum_{i=1}^N \left| \varphi_N\left(\frac{i}{N+1}\right) - \varphi\left(\frac{i}{N+1}\right) \right| \\ & = o(1) N^{-\frac{1}{2}} \sum_{i=1}^N \left| \varphi_N\left(\frac{i}{N+1}\right) - \varphi\left(\frac{i}{N+1}\right) \right| = A_N , \text{ say} \end{aligned}$$

But, $\lim_{N \rightarrow \infty} A_N = 0$, by Puri-Sen (1971), pages 409-411.
 This immediately entails $\frac{S_N - S_N^*}{g} \rightarrow 0$ in probability, which proves Corollary 2.2.

A consequence of Corollary 2.2 is that in many cases of practical interest (such as the normal scores), asymptotic normality holds with the same centering sequence $\tilde{\mu}_N$, whether the scores are given by (1.3) or (1.4). Thus theorem 2.1 and corollary 2.2 serve to unify the results of Hajek (1968) and Hoeffding (1973), and also to some extent simplify the results of the latter paper.

3. Remainder Terms of \tilde{S}_N : In this section, we obtain an estimate for the remainder in the normal approximation to \tilde{S}_N .

Recall that we can write

$$(3.1) \quad \tilde{S}_N - \tilde{\mu}_N = \sum_{i=1}^N \tilde{A}_{Ni} (X_{Ni}) + \tilde{D}_N$$

where

$$(3.2) \quad \tilde{A}_{Ni} (X_{Ni}) = (A_{Ni}^{(1)} (X_{Ni}^{(1)}), \dots, A_{Ni}^{(p)} (X_{Ni}^{(p)}))'$$

$$\tilde{D}_N = (D_N^{(1)}, \dots, D_N^{(p)})'$$

$$D_N^{(\nu)} = D_{1N}^{(\nu)} + D_{2N}^{(\nu)} + D_{3N}^{(\nu)}, \quad 1 \leq \nu \leq p$$

Normalization by s_ν is of no consequence and was introduced there merely to simplify the details of the proof. We shall not

use it here.

We shall simplify the notation somewhat and also express it in a form so that we can use Corollary 17.2, page 165 of Bhattacharya-Ranga Rao (1976).

We can write,

$$(3.3) \quad \tilde{T}_{Ni} = N^{\frac{1}{2}} A_{Ni}(X_{Ni}) , \quad v_i = \text{Cov}(\tilde{T}_{Ni}) .$$

Then,

$$(3.4) \quad \tilde{S}_N = N^{-\frac{1}{2}} \sum_{i=1}^N \tilde{T}_{Ni} + \tilde{D}_N .$$

Let

$$(3.5) \quad v = N^{-\frac{1}{2}} \sum_{i=1}^N v_i .$$

Comparing (3.5) to (2.7) and (2.8), it is clear that the elements of v are $s_{\mu\nu}$.

Let, for $\underline{x} = (x^{(1)}, \dots, x^{(p)})'$,

$$(3.6) \quad F_N(\underline{x}) = P(s_N^{(1)} - \mu_N^{(1)} \leq x^{(1)}, \dots, s_N^{(p)} - \mu_N^{(p)} \leq x^{(p)}) .$$

We then have the following theorem:

Theorem: Let the conditions of Theorem 2.1 be satisfied.

Let in addition,

$$(3.7) \quad \sup_{0 \leq t \leq 1} |\varphi_v'(t)| = \|\varphi_v'\| < +\infty , \quad 1 \leq v \leq p ,$$

and v be positive definite.

Then, there exists a constant $C(p)$, independent of N ,
such that,

$$(3.8) \quad \sup_{\tilde{x} \in \mathbb{R}^p} |F_N(\tilde{x}) - \Phi_{0,v}(\tilde{x})| \leq C(p) \lambda^{-3/2} \rho_3 N^{-\frac{1}{2}} + \Delta_N$$

where

(3.9) $\Phi_{0,v}(\tilde{x})$ is the distribution function (p-variate) of
a Gaussian random vector with mean 0 and dispersion matrix v ;

(3.10) $\lambda = \text{smallest eigenvalue of } v$;

(3.11) $\rho_3 = \frac{1}{N} \sum_{i=1}^N E \|T_{Ni}\|^3$, the norm being the euclidean norm;

and $\Delta_N \rightarrow 0$ pointwise.

Furthermore

$$(3.12) \quad N^{\frac{1}{2}} \Delta_N = O_p \left(\sum_{v=1}^p \|\varphi_v\| s_v^{-1} \left[\sum_{i=1}^N |c_{Ni}^{(v)}| \right] \right)$$

Remark. Since $\Delta_N \rightarrow 0$, it would be of interest to have an estimate for the fluctuations of Δ_N . But in view of the generality of the situation and the fact that we are dealing with the multivariate set up, it is not easy to secure a purely numerical estimate.

(3.12) gives a probabilistic bound. The problem of obtaining a sharp bound remains open.

Proof: Writing $\tilde{T}_N = N^{-\frac{1}{2}} \sum_{i=1}^N T_{Ni}$, we have

$$(3.13) \quad \tilde{s}_N = \tilde{\mu}_N = \tilde{T}_N + \tilde{\Delta}_N.$$

Let G_N be the distribution function (p-variate) of \tilde{T}_N . Then

$$(3.14) \quad F_N(\tilde{x}) = G_N(\tilde{x} - \tilde{\Delta}_N).$$

Hence,

$$(3.15) \quad |F_N(\tilde{x}) - \Phi_{0,v}(\tilde{x})| \leq |G_N(x - \tilde{D}_N) - \Phi_{0,v}(\tilde{x} - \tilde{D}_N)| \\ + |\Phi_{0,v}(\tilde{x} - \tilde{D}_N) - \Phi_{0,v}(x)| .$$

We estimate the two quantities on the right hand side separately.

First, an application of Corollary 17.2 (page 172) of Bhattacharya-Ranga Rao (1976) yields,

$$(3.16) \quad |G_N(x - \tilde{D}_N) - \Phi_{0,v}(x - \tilde{D}_N)| \leq C(p) \lambda^{-3/2} \rho_3 N^{-1/2}$$

uniformly in \tilde{x} , wherein all the quantities have already been defined. $\|\varphi'\| < +\infty$ ensures the existence of ρ_3 .

Estimating the second term is rather more involved. However, since $\Phi_{0,v}$ is continuous, the convergence of the distribution functions is uniform and hence

$$(3.17) \quad \sup_{x \in \mathbb{R}^p} |\Phi_{0,v}(\tilde{x} - \tilde{D}_N) - \Phi_{0,v}(\tilde{x})| = \Delta_N$$

converges to zero pointwise. We only need to obtain (3.12).

Note that, by the Mean Value Theorem (in \mathbb{R}^p), there is a point ξ on the Line segment joining $\tilde{x} - \tilde{D}_N$ and \tilde{x} such that

$$(3.18) \quad \Phi_{0,v}(\tilde{x}) - \Phi_{0,v}(\tilde{x} - \tilde{D}_N) = \frac{\tilde{D}'_N}{\tilde{x}} L(\xi)$$

where $L(\xi)$ is the differential of $\Phi_{0,v}$ at ξ .

Next, for the right hand side of (3.18), we have

$$(3.19) \quad \frac{\tilde{D}'_N}{\tilde{x}} L(\xi) = \sum_{v=1}^p D_N^{(v)} \frac{\delta}{\delta x^{(v)}} |\Phi_{0,v}(\tilde{x})|_{\tilde{x}=\xi}$$

Further, for each v , $1 \leq v \leq p$, it is easy to check that

$\left| \frac{\delta \Phi_{0,v}}{\delta x} \right|$ is dominated by the v^{th} marginal density.

$$\text{Hence, } \left| \frac{\delta \Phi_{0,v}}{\delta x} \right| \leq s_v^{-1} (2\pi)^{-\frac{1}{2}} .$$

Consequently,

$$(3.19) \quad |\Delta_N| \leq \sum_{v=1}^p s_v^{-1} |D_N^{(v)}| (2\pi)^{-\frac{1}{2}} .$$

Thus, (3.12) will be established if we show that for each v ,

$$(3.20) \quad \sqrt{N} D_N^{(v)} = o_p(\|\varphi'_v\| \sum_{i=1}^N |C_{Ni}^{(v)}|) .$$

Again, we shall drop the index v in the rest of the proof.

Rearranging terms of D_N , we get,

$$(3.21) \quad D_N = \int_{-\infty}^{\infty} \left\{ \varphi \left(\frac{N}{N+1} H_N(x) \right) - \varphi(H(x)) \right\} d C_N(x) \\ - \int_{-\infty}^{\infty} (H_N(x) - H(x)) \varphi'(H(x)) d C(x) .$$

In order to simplify the proof, we shall drop the factor $\frac{N}{N+1}$. In view of the fact that φ' is bounded, the conclusion will not be affected.

Consider $\varphi(H_N(x)) - \varphi(H(x))$. By the Mean Value Theorem,

$$\varphi(H_N(x)) - \varphi(H(x)) = (H_N(x) - H(x)) \varphi'(\xi_N(x))$$

for some $\xi_N(x)$.

Hence,

$$(3.22) \quad \left| \sqrt{N} \int_{-\infty}^{\infty} (H_N(x) - H(x)) \varphi'(\xi_N(x)) dC_N(x) \right| \\ \leq \|\varphi'\| \int_{-\infty}^{\infty} \sqrt{N} |H_N(x) - H(x)| d|C_N|(x)$$

Let $\epsilon > 0$ be given. Then by Puri-Sen (1971), there is a constant $C(\epsilon)$ independent of N such that with probability $> 1 - \epsilon$

$$\sqrt{N} |H_N(x) - H(x)| < C(\epsilon) .$$

Thus in (3.22), with probability $> 1 - \epsilon$,

$$(3.23) \quad \left| \int \sqrt{N} (H_N(x) - H(x)) \varphi'(\xi_N(x)) dC_N(x) \right| \leq C(\epsilon) \|\varphi'\| \sum_{i=1}^N |C_{Ni}| .$$

The proof of

$$(3.24) \quad \left| \int_{-\infty}^{\infty} \sqrt{N} (H_N(x) - H(x)) \varphi'(\xi_N(x)) dC(x) \right| \leq C(\epsilon) \|\varphi'\| \sum_{i=1}^N |C_{Ni}|$$

in probability is identical.

(3.23) and (3.24) establish (3.20) and (3.19), which in turn entail (3.12). The proof is completed.

Remark: From the expressions for $s_{\mu\nu}$ and s_{ν}^2 , it is clear that λ is extremely hard to compute. Recall that λ is the smallest eigenvalue of the dispersion matrix V . However, if the components $x_{Ni}^{(\nu)}$ of the vectors \tilde{x}_{Ni} satisfy the

condition of "weak dependence" in the sense

$$(3.25) \quad s_v^2 - \sum_{\mu \neq v} |s_{\mu\mu}| > 0, \quad v = 1, \dots, p;$$

the classical Grischgorin theorem (cf. Dahlquist-Björk, (1974)), permits us to replace λ by much simpler expression (3.25), assuming without loss of generality, $\lambda > 0$. If one is prepared to work numerically, much better estimates for λ can be obtained (op.cit. Dahlquist-Björck).

REFERENCES

- [1] Bergström, H. and Puri, M.L. (1977). Convergence and remainder terms in linear rank statistics. Ann. Stat. 5.
- [2] Bhattacharya, R.N. and Ranga Rao, R. (1976). Normal Approximation and Asymptotic Expansions. John Wiley, New York.
- [3] Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. Ann. Math. Stat. 29, 972-996.
- [4] Dahlquist, G. and Björck, A. (1974). Numerical Methods (English translation by N. Anderson), Prentice-Hall, Englewood Cliffs, New Jersey.
- [5] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. Ann. Math. Stat. 39, 325-346.
- [6] Hoeffding, W. (1973). On the centering of a simple linear rank statistics. Ann. Stat. 1, 56-66.
- [7] Hušková, Marie (1977). Rates of convergence of quadratic rank statistics. J. Mult. Analysis. 7, 63-73.
- [8] Hušková, Marie (1977). The rate of convergence of simple linear rank statistics under hypothesis and alternatives. Ann. Statist. 5, 658-670.
- [9] Jurečkova, J. and Puri, M.L. (1975). Order of normal approximation for rank test statistics distribution. Ann. Prob. 3, 526-533.
- [10] Puri, M.L. and Sen, P.K. (1969). A class of rank order tests for a general linear hypothesis. Ann. Math. Stat. 40, 1325-1343.
- [11] Puri, M.L. and Sen, P.K. (1971). Nonparametric Methods in Multivariate Analysis. John Wiley, New York.

- [12] Puri, M. L. and Rajaram (1977). Asymptotic normality and convergence rates of linear rank statistics under alternatives. Banach Center Publications (to appear).
- [13] van Zuijlen (1977). Empirical distributions and rank statistics. Math. Centre, Amsterdam.

Department of Mathematics
Swain Hall East
Indiana University
Bloomington, Indiana 47401

Department of Mathematics
Merrill Hall
Kent State University
Kent, Ohio 44242